Fermi National Accelerator Laboratory

FERMILAB-CONF-90/172-T
August, 1990

OVERVIEW OF PARTON DISTRIBUTIONS AND THE QCD FRAMEWORK[1]

Wu-Ki Tung

Illinois Institute of Technology, Chicago, Illinois 60616
and
Fermi National Accelerator Laboratory P.O. Box 500, Batavia, Illinois 60510

ABSTRACT

The perturbative QCD framework as the basis of the parton model is reviewed with emphasis on several issues pertinent to next-to-leading order (NLO) applications to a wide range of high energy processes. The current status of leading-order and NLO parton distributions is summarized and evaluated. Relevant issues and open questions for second-generation global analyses are discussed in order to provide an overview of topics to be covered by the Workshop.

1 Introduction

Perturbative Quantum Chromodynamics (QCD) provides the theoretical basis for the intuitively appealing Parton Model. It furnishes a comprehensive framework for describing general high energy processes in current and planned accelerators and colliders. The fundamental formula – the Factorization Theorem – relates a typical *physical cross-section* to a sum over relevant basic partonic *hard cross-sections* (which can be calculated perturbatively) convoluted with corresponding *parton distributions* (which can, in principle, be extracted from a set of standard experiments at moderately high energies). With proper attention to their definition and a consistent convention, these distribution functions are *universal* – *i.e.*, they are independent of the physical process to which they are applied.

In leading-order (LO), the QCD parton framework reproduces original parton model[1] results, with scale-dependent parton distributions. Since the early 70's, this simple model has enjoyed spectacular successes in unifying the phenomenology of all sorts of high energy processes to about the 10–15% level within currently available x range, modulo some overall "K-factors" for processes such as lepton-pair production. Furthermore, it has been used as an indispensable tool to make projections for future physics at much higher energies and small x.[2] In these applications, the well-known leading order parton parametrizations[3,2] played an essential role.

[1]To appear in Proceedings of Workshop on Hadron Structure Functions and Parton Distributions, Fermilab, 1990.

Operated by Universities Research Association Inc. under contract with the United States Department of Energy

In recent years the use of QCD parton model has developed to a stage at which much improved knowledge of parton distribution functions are clearly required along three distinct fronts:

(i) In physical processes which provide precise tests of the Standard Model, the precision of experimental measurements has improved dramatically on the one hand, and the relevant hard cross-sections have been calculated to the next-to-leading order (NLO) and beyond on the other. In order to make real progress, it is crucial to know the parton distributions to a comparable degree of accuracy. This, in turn, prescribes the use of NLO evolution of the distributions and requires a much more detailed comparison with experimental data in the extraction of these distributions, especially for the sea-quarks which are not well determined because of their relative small size. Relevant processes in this category are: electron, muon & neutrino deep inelastic scattering (DIS); lepton-pair production (LPP or DY), and (W, Z, γ) production. It has become increasingly clear that the lack of detailed knowledge of the parton distributions often constitutes the largest source of uncertainty in precise tests of the Standard Model in these processes (e.g. the determination of the Weinberg angle).

(ii) In the study of jet physics and associated production of (W, Z, γ) with jets, which are important on their own right as well as sources of significant background for "new physics", reliable knowledge of the gluon distribution is crucial to the predictive power of the QCD parton model. But the gluon distribution is, so far, not very well determined.

(iii) Physical processes at future colliders (SSC, LHC) involve partons at very small x, well beyond the currently measurable range (around $0.03 < x < 0.75$). In order to make quantitative predictions, it is important: (a) to gain more theoretical insight on the small-x behavior of parton distributions and on the interface of small-x physics to "soft" (e.g. Reggeon) physics; and (b) to determine phenomenologically the range of possible small-x extrapolation of parton distributions consistent with currently available data.

This review obviously cannot cover the entire landscape of activities on Parton Distributions and the Perturbative QCD Framework. In the first part, I shall outline the NLO QCD parton formalism and focus on three examples to illustrate some of its non-trivial features which must be taken into account in any quantitative applications. Although these examples are fairly simple conceptually, not all of them are known to practitioners of QCD phenomenology, resulting in frequent misunderstanding and sometimes misuse of the formalism. In the second part, I shall survey the existing parton distribution parametrizations, summarize the relevant issues confronting "second generation" global analyses of parton distributions, and comment on the proper use of these distributions.

It should be emphasized from the very beginning that the study of parton distributions now encompasses a full range of lepton-hadron and hadron-hadron processes: and, as implied by the above list of motivations, it is inextricably intertwined with all areas of high energy physics from the precision tests of the Standard Model to the search for "new physics". The aim of this review is to help establish a certain common ground and a common language for subsequent discussions in this Workshop, among participants with a diverse background.

2 QCD Parton Model in Next-to-Leading Order

For the sake of simplicity, I shall use a generic lepton-hadron scattering process as the talking point. All the issues I discuss also apply to hadron-hadron scattering, albeit in a somewhat more involved form. The generic process is of the form: $\ell + H \longrightarrow C + X$, where C either represents an identified final-state particle with specific attributes (such as heavy mass or large transverse momentum) or is null (in the case of total inclusive scattering). The "master equation" of the QCD Parton Model is the *factorization* formula [4] which reads:

$$\sigma^i_{H\to C}(q,p) = \sum_a \; f^a_H(\xi,\mu) \otimes \hat{\sigma}^i_{a\to C}(q,k,\mu) \tag{1}$$

where, as illustrated in Fig. 1, H is the target hadron label; a is the parton label; i is the electroweak vector boson helicity label; (q,p,k) are the momenta of the vector boson, the hadron, and the parton respectively; μ is a renormalization scale: and ξ is the fractional momentum carried by the parton with respect to the hadron. The symbol $\otimes$ denotes a convolution (over the variable ξ) of the parton distribution function f^a_H and the hard vector-boson-parton cross-section $\hat{\sigma}^i_{a\to C}$.

The hard cross-section $\hat{\sigma}^i_{a\to C}$ can be calculated in perturbative QCD:

$$\hat{\sigma}^i(\xi, Q/\mu, \alpha_s(\mu)) = \hat{\sigma}^i_0 \; \delta(1-\xi) + \alpha_s(\mu) \; \hat{\sigma}^i_1(\xi, Q/\mu) + O({\alpha_s}^2) \tag{2}$$

where we have suppressed the initial parton label a. The LO $\hat{\sigma}^i_0$ is a constant proportional to the square of the electro-weak coupling of the parton. To calculate the NLO hard cross-section $\hat{\sigma}^i_1$, one encounters divergences which must be subtracted in order to yield finite answers. The subtraction term, in effect, corresponds to that part of the NLO contribution pertaining to almost on-the-mass-shell and collinear parton lines which is already included in the LO term by virtue of the use of QCD-evolved parton distributions. Since the subtraction, hence the hard cross-section $\hat{\sigma}$, is renormalization scheme and renormalization scale *dependent* while the physical cross-section on the left-hand side of Eq. (1) must be *independent* of these theoretical artifices, ***the parton distribution functions f^a_H have to be scheme-dependent objects***

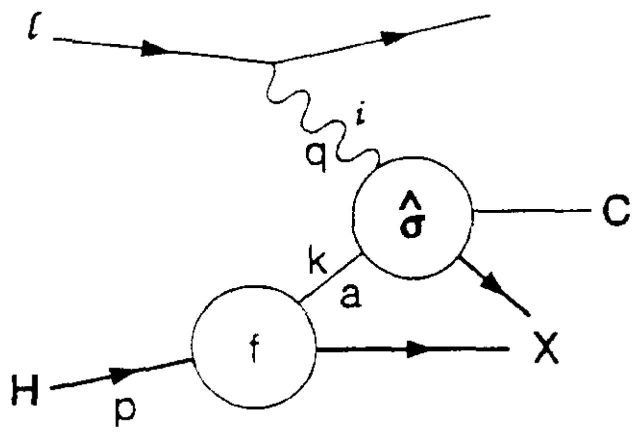

Figure 1: The QCD parton picture and the Factorization Theorem.

to match the definition of $\hat{\sigma}$. We shall demonstrate in Sec. 2.2 that *the scheme-dependence of the gluon and sea-quark distributions can be very substantial* – contrary to conventional expectations. This can lead to important phenomenological consequences. (See also Sec. 2.3)

We begin by examining some basic issues concerning the QCD coupling function α_s in the presence of heavy quarks.

2.1 The QCD Coupling and Λ_{QCD}

The *running coupling* $\alpha_s(\mu)$ is the most basic of QCD quantities. In the well-known case of *all zero-mass quarks*, the standard formulas for α_s in LO and in NLO ($\overline{\text{MS}}$ scheme) are, respectively:

$$\alpha^{LO}(n_f, \mu/\Lambda) = \frac{1}{\beta_0 \log(\mu/\Lambda)} \tag{3}$$

$$\alpha^{NLO}(n_f, \mu/\Lambda) = \frac{1}{\beta_0 \log(\mu/\Lambda)} \left[1 - \frac{\beta_1}{\beta_0^2} \frac{\log\log(\mu/\Lambda)^2}{\log(\mu/\Lambda)^2}\right] \tag{4}$$

where n_f is the number of quark flavors and it enters the right-hand side through the constants,

$$\beta_0 = \frac{33 - 2n_f}{3} \qquad \beta_1 = \frac{153 - 19n_f}{3} \tag{5}$$

If all quarks are massless, the number n_f is fixed and the running coupling α_s is determined by a single parameter Λ – the "QCD lambda".

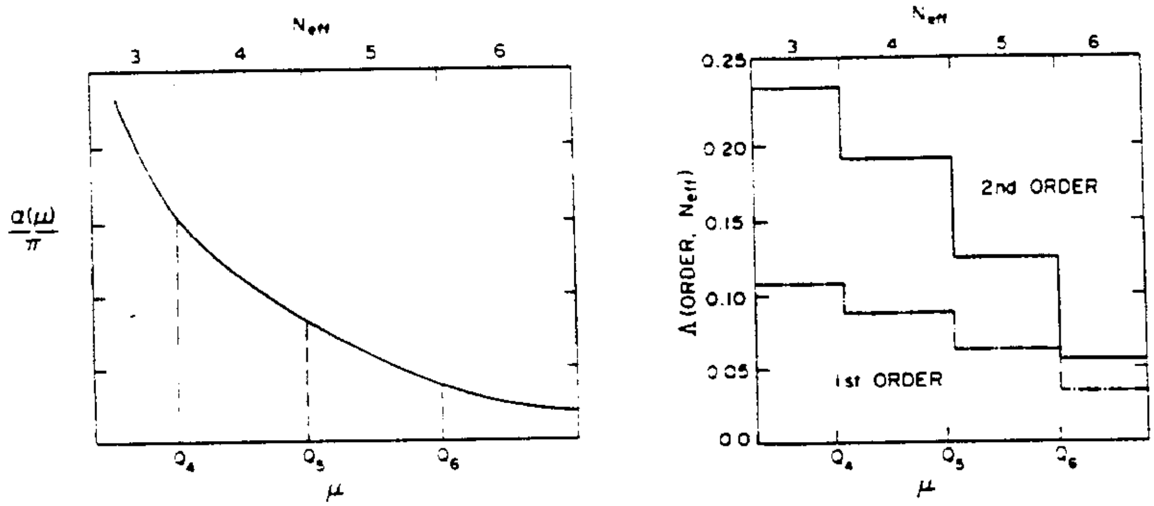

Figure 2: α_s and Λ_{QCD} as functions of μ and n_f^{eff} for a given coupling strength.

In the presence of massive quarks, the situation is quite different. According to the decoupling theorem, each heavy quark i with mass m_i is effectively decoupled from physical cross-sections at energy scales μ below a certain threshold Q_i which is of the order m_i. Thus, *the number of effective quark flavors* n_f^{eff} *is an increasing step function of the scale* μ. Under this circumstance, the specification of the running coupling α_s and the associated Λ_{QCD} is not as simple as before. Although this point is fairly well-known, there still exist considerable confusion and ambiguity about these parameters in both the current literature and in conference presentations. Hence, it is worthwhile to summarize the proper formulation of the problem explicitly.

The definitions of α_s and Λ_{QCD} in the presence of mass thresholds are not unique – they are renormalization-scheme dependent. A *natural* choice is based on the requirement that $\alpha_s(\mu)$ be a continuous function of μ, and that between thresholds, it reduces to the familiar $\overline{\rm MS}$ α_s.[5] This requirement leads to the condition that $Q_i = m_i$ (instead of $2m_i$, or even $4m_i$[2], as often chosen for phenomenological reasons). This choice has the additional desirable feature that the parton distribution functions so defined are guaranteed to be continuous across the thresholds. If Eq. (4) is to remain valid with α_s being continuous in μ, but n_f^{eff} a discontinuous function of μ, it is quite obvious that the effective value of Λ must also make discontinuous jumps with n_f^{eff} at heavy quark thresholds. The same remark applies if one uses the LO formula for α_s, Eq. (3). Fig. 2a shows a typical α_s vs. μ plot; and Fig. 2b shows the corresponding Λ_{QCD} as a function of μ (bottom scale) and n_f^{eff} (top scale).

Fig. 2a explicitly shows that the running coupling function of QCD $\alpha_s(\mu)$ can be unambiguously specified by giving its value at a (standard) scale, say μ_0. On the other hand, as shown in Fig. 2b, this same coupling function is associated with many *different* values of Λ_{QCD}, depending on the number of effective quark flavors and on whether the LO or NLO formula is used. Thus, if one prefers to define α_s by specifying a value of Λ_{QCD}, *it is imperative that one specifies the associated* n_f^{eff} *and the order (LO or NLO) explicitly.* In the recent literature, the second-order $\overline{\rm MS}$ Λ_{QCD} with 4 flavors has increasingly become the standard choice.

2.2 Scheme-Dependence of Parton Distributions

We have mentioned in Sec. 2 that parton distribution functions $f_H^a(x,\mu)$ are renormalization scheme dependent beyond the leading order. In applications to various physical processes, the choice of scheme for the parton distributions must match that of the hard cross-section in the QCD parton model formula.[4] The same parton distribution in two different schemes differ by a well-defined expression which is *nominally* of one order higher in α_s. In this section, we shall point out some non-trivial consequences of the scheme-dependence of parton distributions when the nominal behavior is violated (for a reason). These cases are important in applications, but have not so far received much attention among users of the QCD parton formalism.

To be specific, we consider the often used "DIS" scheme[6] which is defined, in relation to the "universal" $\overline{\rm MS}$ scheme (used by theorists to calculate hard matrix elements), through the following NLO formula for the W_2 structure function of virtual γ deep inelastic scattering (cf. Eq. (1)):

$$\begin{aligned} W_2^{\gamma}(x,Q) &= f^{\rm q}_{\overline{\rm MS}} \otimes \left[C^{(0)}_{2,\rm q} + C^{(1)\overline{\rm MS}}_{2,\rm q} \right] + f^{\rm G}_{\overline{\rm MS}} \otimes C^{(1)\overline{\rm MS}}_{2,\rm G} + O(\alpha_s{}^2) \\ &\equiv f^{\rm q}_{\rm DIS} \otimes C^{(0)}_{2,\rm q} + O(\alpha_s{}^2) \end{aligned} \tag{6}$$

where $C^{(i)}, i = 0,1$ are the hard partonic structure function in LO and NLO ($\hat{\sigma}^i_{a\to C}$ of Eq. (1)), often called the Wilson coefficients in the current context.[7] This formula does not define the gluon distribution in the DIS scheme. It is conventional to require that the momentum sum rule be preserved and then fix the definition of the gluon distribution by generalizing the resulting condition to all moments.[8]

The first point to be made about the DIS scheme is that its definition is designed to render simple the formula for W_2, *and only* W_2. Even in the same scheme, however, the other deep inelastic scattering structure functions – W_1, W_3 or $W_{\rm left}, W_{\rm right}$ – *do* contain non-trivial NLO contributions from both quarks and gluons. It is *not* true, as one might have surmised from the terminology, that all DIS structure functions become simple in the DIS scheme!

In the same vein, the above definition only applies to the *total inclusive* structure function W_2. In practice, one is often interested in semi-inclusive processes such as the production of jets at a given transverse momentum or the production of heavy flavors (charm, bottom, etc.). It would be quite *wrong* to assume *even in deep inelastic lepton-hadron scattering* that, by using the DIS scheme parton distributions, one can neglect NLO terms in the semi-inclusive processes. In fact, the scattering of the electroweak vector boson with gluons in the hadron (a NLO hard process) is *mainly responsible* for producing final state jets at non-vanishing transverse momentum and final state heavy quark flavors (by pair-production), especially at energies not too far above threshold. We will return to this point in the next subsection.

A moment's reflection should also reveal that, although it is theoretically allowed to absorb the entire NLO contribution to W_2 into the definition of the DIS scheme quark distributions (which assumes collinear, on-the-mass-shell partons), this convenience for the total inclusive process comes at the expense of apparent *over-subtraction* (since the NLO diagrams contain non-collinear, off-the-mass-shell quark configurations as well). Thus, in applications to semi-inclusive processes, the use of DIS scheme distributions requires some care, and may lead to counter-intuitive results.

We now show an example of the importance of specifying the scheme in which the parton distributions are defined. Let us consider the question of "hard" *vs.* "soft" gluons which is, so far, an unsettled issue. Without attempting to resolve this problem, we would like to show that its very formulation requires close attention to the defining scheme of the distribution. To wit, the difference between the $\overline{\rm MS}$ and DIS definition of the gluon distribution is (cf. discussion following Eq. (6)):

$$f^{\rm G}_{\overline{\rm MS}}(x,Q) - f^{G}_{DIS}(x,Q) = f^{q_s} \otimes C^{(1)\overline{\rm MS}}_{2,q} + f^{\rm G} \otimes C^{(1)\overline{\rm MS}}_{2,\rm G} \tag{7}$$

where q_s denotes the singlet quark distribution and, in keeping with the perturbative nature of this equation, the scheme label is dropped on the right-hand side. Assuming that one of the f^G's, say f^G_{DIS} is *soft* in one of the schemes – for example, it might behave like $(1-x)^\eta$ with $\eta \geq 6$,[9] then it approaches zero very fast as $x \longrightarrow 1$. However, since the first term on the right-hand side contains a convolution of the *valence quark* distributions with $C^{(1)\overline{\rm MS}}_{2,q}$, it certainly is fairly "hard" – say, behaving like $(1-x)^\eta$ with $\eta \sim 4$.[13] As a consequence, the same distribution in the other scheme ($f^G_{\overline{\rm MS}}$ in this example) will necessarily be hard! Thus *the "hardness" or "softness" of the gluon distribution is a very scheme-dependent concept*, which does not necessarily have an independent meaning. In fact, in the "large x" region (say, $x > 0.4$) where the gluon distribution in the DIS scheme is traditionally considered to be small compared to the valence quark ones, perturbative relations such as Eq. (7) becomes of questionable meaning since a NLO term on the right-hand side becomes larger than one of the LO terms on the left-hand side.

A similar situation exists for the sea-quark distributions *over the entire range of*

x. For this case we have:

$$f_{DIS}^{\bar{q}}(x,Q) - f_{\overline{MS}}^{G}(x,Q) = f^{\bar{q}} \otimes C_{2,q}^{(1)\overline{MS}} + f^{G} \otimes C_{2,G}^{(1)\overline{MS}} \tag{8}$$

It is well-known that, for moderate values of Q, the gluon distribution f^G is much larger numerically than the sea-quark distributions $f^{\bar{q}}$. For instance, in terms of the fractional momentum carried by the partons, the ratio is around 0.50 : 0.03 – a factor of greater than 10. Hence, the second term on the right-hand side of the equation can easily be of the same order of magnitude as the individual terms on the left-hand side in spite of the fact that the Wilson coefficient $C_{2,G}^{(1)\overline{MS}}$ formally carries one extra power of α_s. In other words, *the size of the sea-quark distributions can depend critically on the scheme in which they are defined*; and *it is not very meaningful to talk about a* LO *sea-quark distribution* since its definition is always coupled to the much bigger gluon distribution.[10]

2.3 Order of Magnitude Estimates in QCD

The above observation on the relative order of magnitudes of sea-quarks and gluons is, of course, not surprising, since the sea-quarks are usually understood to arise from the splitting of the gluon. Actually, it is precisely because of this fact that the sea-quark distribution contains an implicit power of α_s with respect to the gluon distribution, at least for moderate values of Q. Thus, excluding the inactive heavy quarks below their thresholds, the parton distributions can be classified into two classes according to their numerical magnitude: (i) those of order 1 – the gluon and the "valence" quarks (u, d) (to be denoted by f^{large} below); and (ii) those effectively of order α_s– the active sea quarks (to be called f^{small}). In this subsection, we discuss important phenomenological consequences of this observation in certain classes of physically interesting processes.

Traditionally, in applying the perturbative QCD formalism to physical processes, the various terms which contribute to the right-hand side of Eq. (1) are classified as LO, NLO, etc., according to the perturbation expansion of the hard cross-section $\hat{\sigma}^{i}_{a\to C}$ only. In view of the large discrepancy in magnitude between the two classes of parton distribution functions f_H^i mentioned above, this traditional power counting can result in misleading conclusions: a "NLO" hard cross-section multiplied by a order-1 parton distribution function can be as important numerically as a "LO" one multiplied by an order-α_s parton distribution function. To make this point explicit, the perturbative QCD formula Eq. (1) can be reorganized, schematically, as follows,

$$\begin{aligned}\sigma_{phys} \; = \; & f^{\mathrm{large}} \otimes \hat{\sigma}^{l}_{LO} + \left[f^{\mathrm{small}} \otimes \hat{\sigma}^{s}_{LO} + f^{\mathrm{large}} \otimes \hat{\sigma}^{l}_{NLO} \right] \\ & + \text{numerically smaller terms}\end{aligned} \tag{9}$$

This point becomes particularly important when the first term on the right-hand

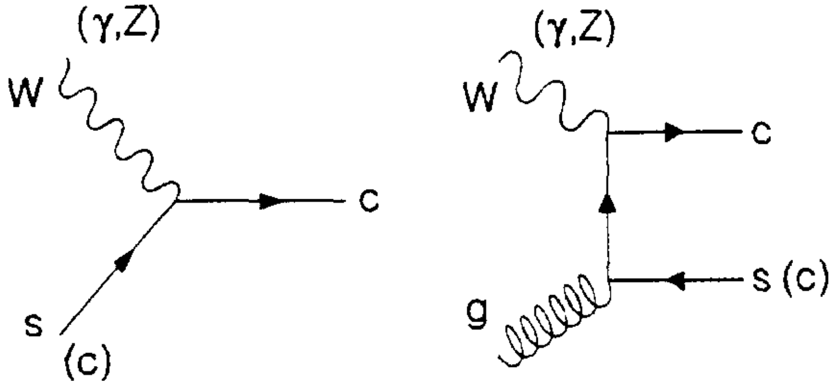

Figure 3: Hard-scattering mechanisms for heavy quark C production in deep inelastic charged-current (neutral-current) scattering.

side is absent or suppressed because $\hat{\sigma}^1_{LO}$ vanishes or contains a suppression factor due to the electro-weak coupling. Then the traditional LO analysis which only keeps the "LO term", $f^{small} \otimes \hat{\sigma}^s_{LO}$, becomes totally inadequate since the "NLO term" in the same square bracket is of the same numerical order. A case in point is charm production in *neutrino deep inelastic scattering. The conventional wisdom is that this* process is dominated by the scattering of the weak W-boson on the strange quark in the hadron target, (Cf. Fig. 3a); and all existing data is analyzed according to this picture. The above discussion clearly suggests that the "NLO" contribution from the scattering of W on the gluons (Fig. 3b) can be just as important, hence must be included in a proper QCD formulation of the problem.[10]

Interestingly, a diametrically opposite approach to this class of problem is also found in the literature, especially involving heavy flavor production in neutral-current processes. It invokes *only* the gluon contribution (the so-called "gluon-fusion" mechanism[11]) (cf. Fig. 3b) and ignores the LO quark diagram Fig. 3a. Although this may make sense in some restricted kinematic region just above the threshold of producing the heavy flavor pair, the LO quark scattering diagram must become increasingly important with increasing energy. The above discussion should make it clear that a consistent treatment must include both mechanisms if it is to be quantitatively reliable over the entire energy range.

We summarize the key point underlying the topics discussed in the last two subsections. Within the QCD parton model, the contributions from the sea quarks and from gluons are always inextricably intertwined. In spite of the conventional designations of LO and NLO respectively, they can be numerically comparable. The precise division between the two mechanisms is tied intimately to the choice of renormalization scheme used during the calculation. Although it is theoretically possible to minimize the contribution from one or the other mechanism to *one given quantity* (e.g. W_2) by a specific choice of scheme. both terms must be included in the analysis of all other physical quantities. In order to achieve consistent results, the choice of scheme must be specified explicitly in these applications.

3 Overview of Global Analyses of Parton Distributions

The global analysis of parton distributions refers to the quantitative comparison of experimental data from a wide range of physical processes with the QCD master equation, Eq. (1), for the purpose of extracting a set of universal parton distribution functions. These can then be used in other applications: to make "predictions" as well as to provide stringent tests of the self-consistency of the perturbative QCD framework itself or of the Standard Model in general. Since any compelling indications of inconsistency of the SM are signs of "new physics", and since even direct search for new physics must rely heavily on understanding of the background from conventional physics, the systematic analysis of parton distributions is intimately tied to all these ventures. To achieve this purpose, contemporary global analyses must incorporate all relevant modern high statistics experimental results and apply the NLO QCD formalism in a consistent manner as described in the first part of this review.

In the following, we: briefly review and assess the existing parton distribution parametrizations (Sec. 3); highlight the relevant phenomenological issues for modern quantitative global analysis (Sec. 3); summarize the current status of on-going programs of global analyses of parton distributions (Sec. 3); and remark on their proper use in phenomenological applications (Sec. 4).

3.1 Review of Parton Parametrizations

The list of widely used parton distributions is a long one. Prominent among these are the pioneering works of Feynman-Field and Buras-Gaemers: followed by the widely used distributions of Gluck-Hoffmann-Reya, Duke-Owens, and Eichten-Hinchliffe-Lane-Quigg.[2,3] These are all based on leading order QCD-evolved distributions extracted by comparison with data existing up to about 1983. In recent years, many second generation high statistics experiments have become available and more refined global analyses have been carried out in NLO by Martin-Roberts-Stirling[12], Diemoz-Ferroni-Longo-Martinelli[8], Aurenche-Baier-Fontannaz-Owens-Werlen[13], and Morfin-Tung[14], reflecting the needs of the present time. Table I lists most of the currently used parton distributions and the experimental data on which the analyses were based.

	D-O[3]	EHLQ[2]	(H)MRS[12]	DFLM[8]	ABFOW[13]	M-T[14]
ν-DIS	CDHS	CDHS	CDHSW (CCFRR)	CHARM	—	CDHSW (CCFR)
μ-DIS	EMC	—	EMC.BCDMS	—	BCDMS	EMC,BCDMS
D-Y	E288,ISR	—	(E288),(E605)	—	—	E288,E605
Dir-γ	—	—	WA70	—	WA70	—

Table I: Parton Distribution sets and data used.
References to the experiments are: BCDMS,[15] CDHS,[16] CDHSW,[17] CCFR,[18] CCFRR,[19] E288,[20] E605,[21] EMC,[22] WA70.[23] Parentheses around entries indicate that the corresponding data were only used partially.

The experimental developments which have the most significant impact on recent analyses as compared to the previous LO ones are: (i) results of the high statistics CDHS[16] neutrino experiment, on which most earlier analyses were heavily dependent, has since been considerably revised and supplanted by the new CDHSW[17] results; (ii) the very accurate new data on muon scattering from the BCDMS[15] collaboration does not fully agree with earlier results, especially the previous "standard" EMC.[22] Both of these developments cause the predictions of the earlier parton distributions on deep inelastic scattering – the main source of information on these distributions – to disagree with the best current data up to 15-20%. Differences of this size correspond to many standard deviations in these high statistics experiments.

The disagreement between the BCDMS and EMC results has been a source of much uncertainty and discussion for the past two years. However, recent comprehensive studies,[24] including the introduction of the new analysis of the SLAC-MIT experiments[25] as an independent check on the normalization, indicate possible ways to resolve the discrepancy. Several talks in this Workshop, including the first report on a new analysis of the EMC data[26], will shed further light on this issue.

To illustrate the general state of affairs, we show one plot on the comparison of current data with the results calculated from representative parton distribution sets in Fig. 4. This is typical of similar plots on the comparison of BCDMS and CDHSW data with the same parton distributions. Details can be found in some recent reviews.[27][28] Fig. 4 clearly shows the necessity for using up-to-date parton distributions which take into account of all relevant experiments in any QCD analysis where precision is required.

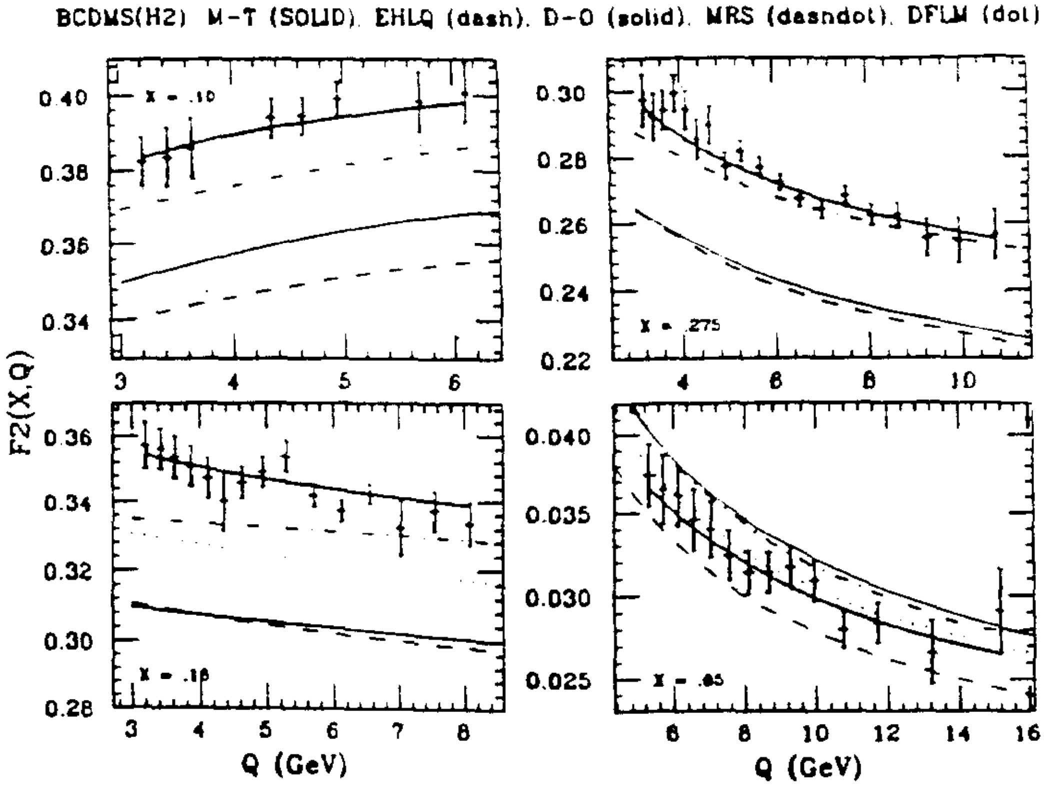

Figure 4: Comparison of a subset of BCDMS F_2 data[15] on hydrogen with calculated results from representative parton distribution sets.

3.2 Relevant Issues for Quantitative Global Analysis

A truly quantitative global QCD analysis of data involves many experimental and theoretical considerations which may affect the results, not all of which are generally known. We shall briefly summarize the important ones. A systematic discussion can be found in the report of the Structure Function and Parton Distribution group in the 1988 Snowmass Proceedings.[27]

On the experimental side, in addition to the choice of physical processes and experiments to fit, all the following factors can affect the consistency and the correctness of the results: (i) the selection of data within a given experiment, according to kinematic cuts in Q^2, W or other variables. (How do the results depend on the values of these cuts?) (ii) Are systematic errors included in the fitting procedure? This is, in fact, a critical issue for the second generation analysis since: (a) errors on current high statistics experimental data are dominated by systematic errors; and (b) when data from several experiments are used in a chi-square or likelihood analysis, the fitting procedure is simply meaningless without including the systematic errors.

However, most existing global analysis *do not* include systematic errors.[1] (iii) Do the different experimental data sets apply the same "corrections" (e.g. "slow-rescaling", "isoscalar", etc.) to their data analyses; and, if not, how should one handle the differences?[27] Unfortunately, real differences can often be found in published data among different experiments, and they are usually overlooked. (iv) Finally, distinct physical quantities measured in the same experiments can have correlated errors. A proper fit to the data must take into account the correlations. No global analysis to date has attempted to incorporate this in a systematic way.

On the theoretical side, the parton distributions and QCD parameters one obtains from global fits can depend (in addition to the choice of LO or NLO formalism) on: (i) the functional form used for the initial distributions, especially if it happens to be too restrictive; (ii) the number of parameters which are allowed to vary when the fit to data is made. Unfortunately, there is no proven way to assure a correct choice on either of these considerations. A reasonable choice when fitting a given set of data may not remain so when additional experiments and/or physical processes are incorporated. Finally, results on these global analyses in the very small x region – a region of much interest in applications to "predict" the high energy behavior of standard and new physics – are heavily dependent on whether the small-x *extrapolation* is fixed by an assumed functional form or is characterized by parameters to be fit to existing data.

These issues, as summarized above, should not only be of concern to those working on global fits; the users of parton distributions must be aware of these considerations and their impact on the physics they are trying to extract from the physically measured quantities.

3.3 Current Status of Global Analyses and Open Issues

In this Workshop we shall hear progress reports on two of the currently active programs on global analysis of parton distributions.[12,14][29][30] It suffices to say that in spite of past efforts and the current detailed work, there are still many open questions which will require a combined effort of refined and expanded experiments, further theoretical clarifications, and continued phenomenological analysis to be resolved. The list includes:

(i) The gluon distribution: What is the proper or optimal definition? (Cf. Sec. 2.2 and [31]) How can it be determined in an unambiguous way? DIS without a reliable measurement of the longitudinal structure function does not determine the gluon

[1] This is ironic, as most experiments devote more effort on understanding the systematic errors than on anything else.

distribution well. Lepton-pair production data imposes better constraints (through combined effects of sea-quarks and gluons.) The "dedicated study" of the gluon distribution based on direct-photon production[13] show a great deal of promise, but is also subject to a number of experimental limitations (limited kinematic range. large errors, isolation criteria, etc.) and theoretical uncertainties (choice of scales, bremsstrahlung contributions, etc.). Can these be satisfactorily resolved? This will be discussed at this Workshop. Are there other comparable or better ways to determine the gluon distribution?

(ii) The sea-quark distributions: How should they be defined? (Cf. Sec. 2.2 and Sec. 2.3) What is the dependence of $f^{q^i}_{sea}$ on the flavor label i? Is it SU(3)-symmetric, SU(2)-symmetric, or asymmetric? Current data on charm production (dimuon final state) in DIS have been interpreted, in the LO picture, to indicate a non-SU(3)-symmetric sea.[32] However, the discussion of Sec. 2.3 suggests that this interpretation needs to be reassessed because of the inherent sear-quark – gluon mixing.[10]

(iii) Small-x and large-x behavior of the parton distributions: In the familiar perturbative QCD formalism, the Q-dependence of parton distributions $f^i(x, Q)$ is governed by the evolution equation with calculable kernels; however, the x-dependence beyond currently measurable range is unknown except for some qualitative guidelines based on Regge-type of arguments at some unspecified scale. Much attention has been given to developing theoretical tools to extend our understanding of the parton model into the small-x region.[33] There are several distinct aspects to the "small-x problem": (i) the resummation of large $(\alpha_s \log(1/x))^n$ terms for fixed Q; (ii) the region of large $\log(1/x)$ *and* large $\log Q$; and (iii) the region of saturation of parton densities and the breakdown of the parton picture as we know it. Promising recent progress will be presented in this Workshop by Levin.[34]

From the phenomenological point of view, one can ask: what reasonable constraints on the small-x extrapolation of the parton distributions can be obtained by global fits to current data? Shown in Fig. 5 are plots of the gluon distribution and the predicted structure function F_2 extending to very small x from two sets of NLO parton distributions which both fit current data well.[14] The difference between the extrapolated results is seen to be quite large. This difference can be resolved by direct measurement of F_2 (and hopefully $G(x, Q)$ as well) at HERA as well as by some suitable hadron-hadron collision process such as lepton-pair production (DY) at the colliders. Fig. 6 illustrates this point by showing the anticipated rapidity distribution of the lepton-pair at the Tevatron for $Q = 20GeV$ using several choices of parton distributions, including the two mentioned above.

The behavior of parton distributions near $x \sim 1$ has also been of interest. The experimental consequences of this are concentrated in the relatively low energy region

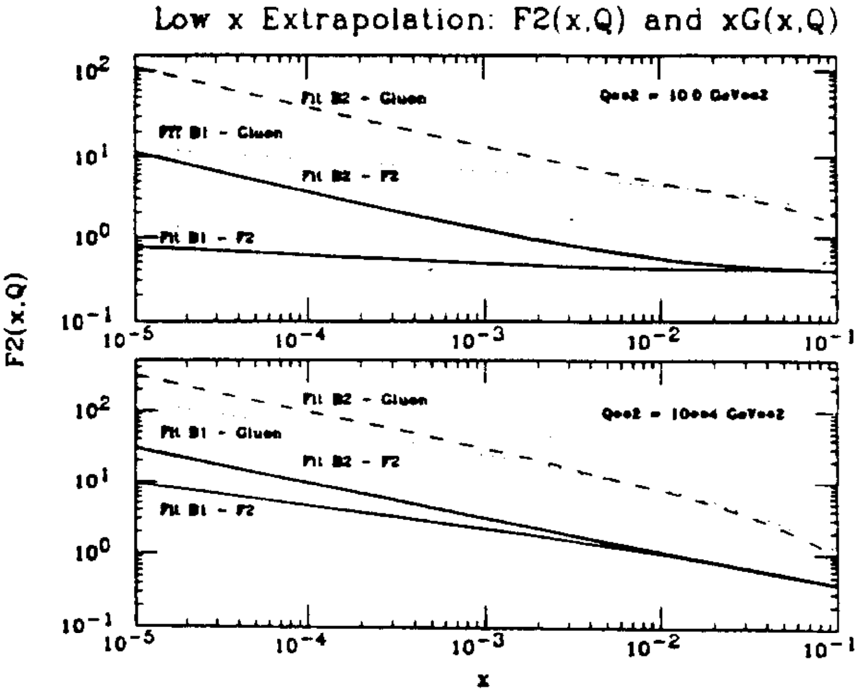

Figure 5: Small-x extrapolation of $G(x,Q)$ and $F_2(x,Q)$ based on two global fits to current data.[14]

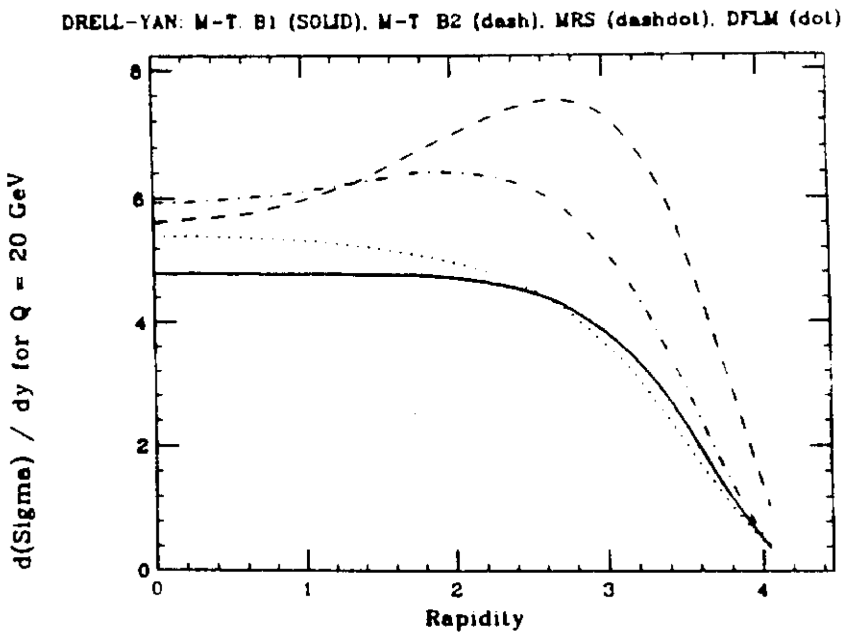

Figure 6: Rapidity distribution of Drell-Yan pairs predicted by some representative parton distributions.[14]

where they are also related to *higher twist* corrections to the conventional leading power-law QCD results.[35]

In order to address the open questions described above, we need to go beyond the traditional reliance on DIS and lepton-pair production processes. While important progress will continue to be made in these areas, especially with the exciting expansion of experimental range offered by HERA, the coming of age of fixed-target direct photon experiments and quantitative measurements of an ever wider range of processes at the hadron colliders have opened up many more possibilities of determining the remaining uncertainties of parton distributions and testing the consistency of QCD. Foremost among these are W-, Z-production cross-sections (including rapidity and transverse momentum distributions), heavy flavor production, jet production (with or without associated vector bosons), as well as DY and direct-photon processes. The program of this Workshop in the coming days underlines this expanding horizon. As most of this work in has yet to be carried out, the quantitative analysis of parton distributions will have to be an on-going effort for some time to come.

4 Some Remarks on the Use of Parton Distributions

As the parton model advances from the original "naive" genre to the "QCD-improved" kind summarized here, its use requires certain reassessment. Here we touch upon some of these considerations, based on discussions of the previous sections.

The LO QCD parton model is still a very powerful and simple framework which describes a wide range of physical processes to within 10–20% accuracy. In these applications, it is sufficient to use LO parton distributions. In fact, it is preferable to do so, than to use the NLO distributions – as the latter are extracted by comparing complete NLO formulas with experiment. (A built-in error is always incurred by the mixed use of a LO hard cross-section with NLO parton distributions.) It is important, however, that the LO distributions used accurately reflect existing data where applicable. As discussed in Sec. 3.1, the most often used first generation parton distributions, unfortunately, strongly disagree with the recent high precision DIS and DY data.

For applications which require more accuracy, the LO formalism is known to be inadequate. When NLO hard cross-section formulas are used in these more refined cases, it makes no sense to use LO parton distributions, since the supposed extra accuracy gained in the use of the former is totally lost by the use of the latter (especially, in addition, if the latter does not fit known current data). Obvious and commonsensical as this remark may appear, the common practice of using well-known distributions under all circumstances in the "comparison of experiments with QCD" is only too apparent in the literature.

A common *myth* in contemporary phenomenology is to apply a variety of parton distributions to a given physical process and then cite the range of results obtained as the "theoretical error". There is no justification for such a practice, since: (i) the LO and NLO distributions are not designed to be used the same way; and (ii) some of these distributions are already known to disagree with current data, as mentioned above. In fact, as shown in Fig. 4, in many cases the deviations from the correct results are all to the *same* direction for most of these distributions rather than "bracketing" the right answer. Since existing data do not completely determine all the parton distributions, it is, of course, useful to explore the range of uncertainties due to this lack of knowledge both in performing precision tests and in making predictions for the future. One can do this properly by obtaining a range of parton distributions in the same global analysis with allowance made for the uncertain feature to be studied. This approach has been used in characterizing the small-x extrapolation of parton distributions in recent studies.[12,14] It can also be applied to other features such as the flavor dependence of the sea quarks.[14]

In the use of NLO QCD formalism, it is important to make consistent choice of renormalization scheme in the hard cross-section formula and in the definition of the parton distributions used (cf. Sec. 2.2). Although the outcome of a given global analysis of parton distributions can, in principle, be presented in any scheme, most authors publish their results only in the scheme which was used in the original analysis. It is, therefore, incumbent on the user to make the necessary adjustment in order to bring about consistency. It is helpful to know that among the published NLO parton distributions, the MRS[12] and ABFOW[13] ones are in the $\overline{\mathrm{MS}}$ scheme, whereas the DFLM[8] ones are in the DIS scheme. The new MT[14] distributions are given in both schemes.

All QCD "predictions" are subject to an uncertainty associated with the choice of renormalization and factorization scales. This uncertainty diminishes at very high energies or, in principle, when more and more higher order terms are included. At current energies and to NLO only, this scale-dependence can be substantial for certain processes. There is, so far, no clear consensus among theorists on whether there is some sensible method for making an intelligent (if not "correct") choice of scale for a given situation. Because of the practical importance of this issue, there will be continued discussion and debate about this topic in this Workshop and beyond.

Finally, the various current parton distribution sets differ considerably in: (i) the choice of input experimental data (cf. Table I); (ii) the treatment of experimental errors and corrections (cf. Sec. 3.2); and (iii) the selection of functional forms for the input distributions and the number of free parameters (cf. Sec. 3.2). Many implicit assumptions ride on the choice made in (iii) which users do not see explicitly. For critical applications, these factors must be examined carefully before definite conclusions are drawn.

5 Summary

Perturbative QCD stands as one of the main pillars of contemporary high energy physics. It plays a central role in pushing quantitative tests of the Standard Model to ever increasing accuracy (hopefully, even to its eventual limit); in studying the signals and backgrounds for new physics; and in providing hints for attempts to formulate non-perturbative solutions to strong interaction physics. Most applications of QCD are based on the Factorization Theorems which have a simple parton model interpretation. They rely on the use of the universal parton distributions to relate measurable quantities to the fundamental processes of the underlying theory.

This Workshop is concerned with a critical review of traditional approaches to the study of parton distribution functions (by DIS, DY, etc.) and, more importantly, a survey of the expanding scope of all the fields which contribute to and make essential use of the QCD parton model. This brief introductory talk underlines some of the important considerations which must be taken into account in applying this general framework to the expanded horizon; and highlights the current status and the open questions of the global analysis of parton distributions, in the hope of helping to establish a useful starting point for the specific sessions to follow in the next three days.

We have confined our survey mainly to NLO perturbative QCD. The thrust of most theoretical studies relating to this field is, of course, concerned with pushing the frontier beyond NLO QCD. Thus, in addition to reports on recent work on small-x and higher twist mentioned in Sec. 3.3, there will be talks on the resummation of large hard cross-sections[36,31], and on multi-parton processes.[37] Finally, there will be a session on spin-dependent parton distributions which pose some unique theoretical and experimental challenges which we have not had time to include in this talk.

Acknowledgement

Many of the issues described in this survey crystalized during discussions I have had with my collaborators M. Aivazis, J. Morfin, F. Olness and numerous other colleagues, too many to mention individually. Particular thanks are due, however,

to John Collins for sharing his unusual insight in perturbative QCD which underlies much of this talk.

References

[1] R.P. Feynman, *Photon-Hadron Interactions*, Benjamin, (1972).

[2] E. Eichten *et al*, *Rev. Mod. Phys.*, **56**, 579 (1984) and Erratum **58** 1065 (1986).

[3] M. Glueck *et al*, *Z. Phys.* **C13**, 119 (1982); D. Duke and J. Owens, *Phys. Rev.* **D30**, 49 (1984).

[4] For a recent review and original references, see J. Collins, D. Soper and G. Sterman in *Perturbative Quantum Chromodynamics*, A. Mueller, Ed., World Scientific Pub., 1989.

[5] J.C. Collins, F. Wilczek, and A. Zee, *Phys. Rev.* **D18**, 242 (1978); J.C. Collins and Wu-Ki Tung, *Nucl. Phys.* **B278**, 934 (1986); See also W. Marciano, *Phys. Rev.* **D29**, 580 (1984).

[6] G. Altarelli, R.K.Ellis, & G.Martinelli, *Nucl. Phys.* **B143**, 521 (1978); and *ibid.* **B157**, 461 (1979).

[7] W. Furmanski & R.Petronzio, *Z. Phys.* **C11**, 293 (1982).

[8] M. Diemoz *et al*, *Z. Phys.* **C39**, 21 (1988).

[9] See. *e.g.* Eichten *et al*, Ref.2; Diemoz *et al*. Ref.8.

[10] M.A.G. Aivazis, F. Olness & Wu-Ki Tung, Fermilab-Pub-90/23 and IIT-90/10.

[11] J. Leveille, T. Weiler, *Nucl. Phys.* **B147**, 147 (1979); T. Weiler, *Phys. Rev. Lett.* **44**, 304 (1980); M. Arneodo, *et al*, *Z. Phys.* **C35**, 1 (1987).

[12] A.D. Martin, R.G. Roberts & W.J. Stirling, *Phys. Rev.* **D37**, 1161 (1988), *Mod. Phys. Lett.* **A4**, 1135 (1989); P.N. Harriman, A.D. Martin, R.G. Roberts & W.J. Stirling, Rutherford Lab. preprint RAL-90-007 (1990); RAL-90-018 (1990).

[13] P. Aurenche *et al*. *Phys. Rev.* **D39**, 3275 (1989).

[14] J.G. Morfin and Wu-Ki Tung, Preprint Fermilab-Pub-90/24, IIT-90-11.

[15] A.C. Benvenuti*et al*, *Phys. Lett.* **B223**, 485 (1989) and CERN-EP/89- 170,171, December, 1989.

[16] H. Abramowicz *et al*, *Z. Phys.* **C17**, 283 (1984); *Z. Phys.* **C25**, 29 (1984); *Z. Phys.* **C35**, 443 (1984).

[17] J.P.Berge *et al*, Preprint CERN-EP/89-103 (1989).

[18] Private communications from Sanjib Mishra.

[19] D.B. MacFarlane *et al*, *Z. Phys.* **C26**, 1 (1984).

[20] A.S.Ito *et al*, *Phys. Rev.* **D23**, 604 (1981).

[21] C.N. Brown *et al*, *Phys. Rev.* Lett. **63**, 2637 (1989).

[22] J.J. Aubert *et al*, *Nucl. Phys.* **B293**, 740 (1987).

[23] M. Bonesini *et al*, *Z. Phys.* **C38**, 371 (1988).

[24] See, e.g. J. Feltesse, *Proceedings of the XIV International Symposium on Lepton and Photon Interactions*, Stanford, August 1989.

[25] A. Bodek, in these Proceedings.

[26] S. Wimpenny, in these Proceedings.

[27] Wu-Ki Tung *et al*, in *Proceeding of the 1988 Summer Study on High Energy Physics in the 1990's*, S. Jensen, Ed., World Scientific, (1990).

[28] K. Charchula *et al*, DESY report DESY 90-019 (1990).

[29] See A. Martin and R. Roberts, these Proceedings.

[30] J. Morfin, these Proceedings.

[31] K. Ellis, these Proceedings.

[32] CDHS: H. Abramowicz, *et al*, *Phys. Rev. Lett.* **57**, 298 (1986); *Z. Phys.*, **C28**, 51 (1985); CHARM: J.V. Allaby, *et al*, *Z. Phys.* **C36**, 611 (1987); CCFR: K. Lang *et al*, *Z. Phys.* **C33**, 483 (1987); S.R. Mishra *et al*, in *Proceedings of 14th Rencontres de Moriond*, Mar. 1889.

[33] For a comprehensive review of pioneering works in this field, see: L. Gribov, E. Levin, and M. Ryskin, *Phys. Rep.* **100**, 1 (1983).

[34] E. Levin, these Proceedings; See also Proceedings of a forthcoming *Workshop on Parton Distributions at Small-x*, DESY, May, 1990.

[35] J.W. Qiu, these Proceedings.

[36] G. Sterman, these Proceedings.

[37] D. Treleani, these Proceedings.